Solution Brief

High-Performance Disaggregated Storage Platform with LeilFS and OpenFlex® Data24 4000

Challenges

- Rapid growth of unstructured data (AI/ML, analytics, media, backup) requires scalable storage architectures.
- High latency in traditional shared storage affecting performance-sensitive workloads.
- Inefficient utilization of storage in siloed or over-provisioned infrastructures.
- Complexity in scaling storage independently from compute resources.
- Increasing power, cooling, and operational costs in large-scale deployments.
- Limited flexibility in legacy architectures (tight coupling of compute and storage).
- Data resiliency, availability, and integrity challenges in distributed environments.
- Need for seamless expansion without downtime or performance degradation.
- Integration of heterogeneous storage platforms and protocols within a unified data access framework.

Highlights

- Distributed POSIX-compliant file system (LeilFS) for scalable and reliable data access.
- NVMe-over-Fabrics (NVMe-oF) powered disaggregated storage using OpenFlex Data24.
- High availability and fault tolerance through distributed architecture with EC and Replication.
- Support for flexible scaling-compute and storage independently.
- Open composable infrastructure enabling dynamic resource allocation.
- Efficient storage utilization with shared NVMe storage.
- Enterprise-grade performance, reliability, and efficiency.
- Reduced operational complexity.
- Lower risk during future platform modernization initiatives.

The integration of LeilFS, a distributed and scalable POSIX file system, with WD OpenFlex Data24 NVMe-oF™ storage platforms enables a modern, high-performance storage architecture. This solution decouples storage from compute, allowing multiple hosts to access shared NVMe™ flash storage with near-local latency while maintaining the scalability and flexibility required for modern workloads.

LeilFS provides the software-defined storage layer, handling data distribution, metadata management, redundancy, and fault tolerance. OpenFlex Data24 delivers high-performance NVMe flash storage over Ethernet fabric using NVMe-oF, ensuring ultra-low latency and high throughput. Together, they form a composable, scalable, and efficient storage solution suited for AI, HPC, cloud, NeoCloud, and enterprise workloads.

Key Solution Components

LeilFS Distributed File System

- Metadata servers, chunk servers, and client architecture
- High availability via replication and erasure coding
- Parallel I/O for high throughput

OpenFlex Data24 NVMe-oF Platform

- 24-bay NVMe SSD enclosure
- 100GbE connectivity with NVMe-over-Fabrics
- Shared storage across multiple hosts
- High-Speed Fabric Network with RoCE or TCP connectivity for flexible deployment
- Composable Infrastructure Layer with Dynamic storage allocation across workloads

OpenFlex Data24 4000 Series Storage Platforms

The OpenFlex Data24 4000 series NVMe-oF storage platform extends the high performance of NVMe flash to shared storage. The 4000 series provide low-latency sharing of NVMe SSDs over a high-performance Ethernet fabric to deliver similar performance to locally attached NVMe SSDs. WD RapidFlex™ NVMe-oF controllers allow up to six dual pathed hosts to be attached without a switch. The OpenFlex Data24 4200 series uses WD's RapidFlex C2000 Fabric Bridge Adapters to provide up to 12 ports of 100GbE which can connect to RDMA and/or TCP configured host ports.

High-Performance Disaggregated Storage Platform with LeilFS and OpenFlex Data24 4000

Features

- NVMe-oF architecture
- Bandwidth match between PCIe® Gen4 SSDs (Storage) and I/O (Network) – No oversubscription
- RESTful API support for simplified management
- N+2 fan redundancy eliminates the need for hot- swappable fan modules Data24 4100
- Connect directly to an Ethernet NIC in your host or through an Ethernet switch
- Device Sharing optional configuration utilizes NVMe Non-Transparent Bridging to allow all Ethernet ports to access any SSD
- High IOPS and throughput performance
- Software-defined composability
- Supports Ethernet-based disaggregated storage
- High availability with redundant components
- Ideal for AI/ML, real-time analytics, and transactional workload

LeilFS

LeilFS is a distributed, open-source POSIX file system inspired by Google File System, designed for high scalability, performance, and reliability across commodity hardware environments. It separates metadata and data paths, enabling parallel data access and efficient scaling while maintaining data integrity and fault tolerance through replication or erasure coding. It employs a chunk-based storage architecture, segmenting files into 64 MiB chunks subdivided into 64 KiB blocks, each with 4 bytes of CRC (Cyclic Redundancy Check) for data integrity.

The replication goal for specific files and directories can be configured to use Erasure Coding based on Reed-Solomon for redundancy purposes. For instance, a file with 64 MiB of data, with replication goal EC(4+2), will divide the chunk into 4 data parts of 16 MiB and 2 parity parts of 16 MiB. This way, there is no data loss even if two Chunk servers are down or any two parts are lost. The system also prioritizes data resiliency through data scrubbing and CRC32 checksum verification. Additional features include instant copy-on-write snapshots, efficient metadata logging, and hardware integration without downtime.

Key Capabilities

- Distributed architecture with metadata servers and chunk servers
- High availability using failover mechanisms
- Data integrity through CRC validation and scrubbing
- Seamless scalability—add nodes without downtime
- Support for petabyte to exabyte-scale deployments
- POSIX compatibility for easy integration with existing applications Up to 60 Ultrastar HDDs (SAS or SATA)

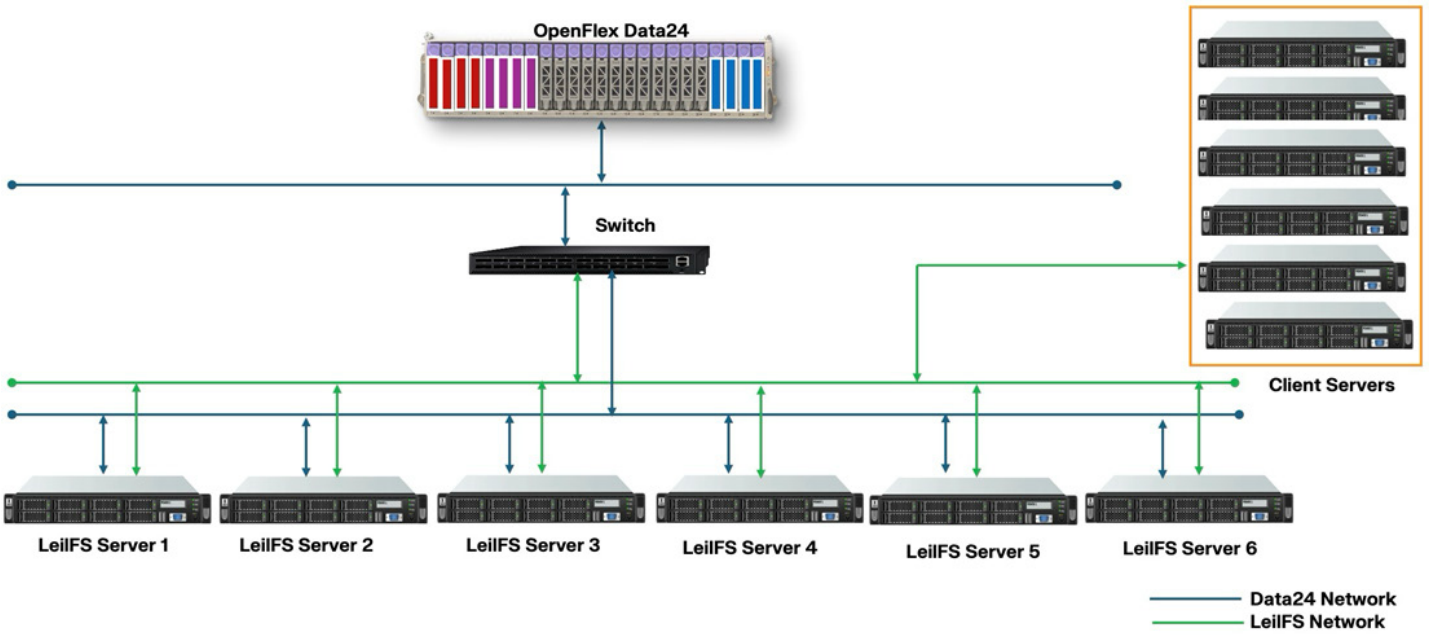
Reference Architecture

In the reference architecture, LeilFS integrates seamlessly with OpenFlex Data24 4000 Series platforms to create a high-performance, disaggregated storage environment where compute, storage, and networking are decoupled yet tightly coordinated through a high-speed fabric. At the core of this design, application or compute nodes access storage through a 100GbE NVMe-oF network, which enables multiple hosts to share NVMe SSD resources with latency and performance comparable to locally attached drives. The OpenFlex Data24 acts as a shared NVMe storage, exposing up to 24 NVMe SSDs through high-bandwidth Ethernet interfaces, eliminating traditional storage silos and ensuring balanced bandwidth without oversubscription across the system.

On top of this high-performance storage layer, LeilFS provides the distributed file system intelligence, organizing storage into a scalable architecture consisting of metadata servers (master, shadow, metaloggers) and data servers (chunkservers), enabling efficient data distribution and parallel I/O operations. Clients interact with metadata servers to determine data locations while simultaneously accessing chunkservers directly for high-throughput data transfer, ensuring minimal bottlenecks. This parallel data path, combined with NVMe-oF's low-latency fabric, delivers a highly optimized environment for data-intensive workloads.

The architecture also ensures high availability and resilience through LeilFS mechanisms such as replication, erasure coding, and failover, while OpenFlex provides hardware-level redundancy (dual controllers, power supplies, and network paths). Together, they enable seamless scalability, allowing additional compute nodes or storage capacity to be added without disrupting existing operations. This composable setup supports dynamic allocation of storage resources to workloads such as AI/ML, HPC, analytics, and enterprise file services, making the reference architecture both flexible and future-ready.

High-Performance Disaggregated Storage Platform with LeilFS and OpenFlex Data24 4000



Performance Details

The measured performance numbers are specific for the current configuration of LeilFS with OpenFlex Data24 in conjunction with the uses of 24 NVMe SSDs with multipath connection via 6 storage nodes. FIO tool is used to run the performance benchmarking on the system. The performance test is run using FIO, simultaneously running tests from all the 6 Ubuntu clients.

LeilFS Performance Details on OpenFlex Data24 4000

Policy Details	Read Throughput	Write Throughput
Goal1 - No Redundancy	40.07 GBps	35.97 GBps
EC (4+2)	36.70 GBps	22.02 GBps
Goal3 - (3* Replication)	39.87 GBps	12.51 GBps

Notes:

- The performance results are captured after running multiple iterations with different data size (file size, block size, concurrency, policies).
- LeilFS uses goals to define filesystem policies. Goal1 represents a filesystem configuration without any redundancy, while Goal3 represents a filesystem configured with a 3-way replication policy.

Disclaimer: All the performance data are subject to change depending on the Drive model, capacity, workloads, servers, CPU, tunings and topology used.

Configurations Details[1]

Metrics	Details
Storage Product	OpenFlex Data24 4213 , Bundle FW : 52 , 3.0.1.35
LeilFS Version	LeilFS 5.10.0-official
Storage Interface	12x QSFP28 Interface (12 x 100Gbps per port)
Server Details	PowerEdge® R550
Host OS	Ubuntu 24.04.3 LTS
Host NIC	CX6
Firmware Version	24.10-5.1.6 (22.43.8016 (MT_0000000437))
CPU	Intel® Xeon® Gold 5318Y CPU @ 2.10GHz or Higher CPU Configuration
CPU Core Details	Dual socket server with 24 core CPU each. 96 logical cores in total with HT enabled
Memory	256 GB
Switch Details	NVIDIA® Onyx™ 3700
Drive Details	Sandisk® SN655 (3.84TB)[2], WUS5EA138ESP7E4, RC610008
Tool	fio-3.36

[1] Actual result depend on CPU memory bandwidth, RNIC count/link speed, PCIe layout, and fabric configuration.
[2] One terabyte (TB) is equal to one trillion bytes and one petabyte (PB) is equal to 1,000 TB. Actual user capacity may be less due to operating environment.

High-Performance Disaggregated Storage Platform with LeilFS and OpenFlex Data24 4000

Advantages

- **High Performance at Scale**: Combines NVMe-oF powered OpenFlex Data24 with LeilFS's parallel I/O design to deliver near-local NVMe performance across distributed environments.
- **Distributed Architecture Efficiency**: LeilFS separates metadata and data paths, enabling clients to access data directly from storage nodes in parallel, reducing bottlenecks and improving throughput.
- **Scalability Without Disruption**: Seamlessly scale storage and compute independently by adding nodes or NVMe capacity without impacting ongoing operations.
- **High Availability & Fault Tolerance**: LeilFS provides built-in redundancy through replication and erasure coding, while OpenFlex ensures hardware-level resilience.
- **POSIX Compatibility**: Full POSIX compliance ensures easy integration with existing enterprise applications and workflows.
- **Optimized Resource Utilization**: Shared NVMe storage pools eliminate silos and reduce overprovisioning, improving infrastructure efficiency.
- **Data Integrity & Reliability**: End-to-end data integrity with CRC verification and periodic validation ensures consistent and reliable storage operations.
- **Composable Infrastructure**: OpenFlex Data24 enables dynamic allocation of storage resources, while LeilFS ensures intelligent data placement and access.
- **Flexible Protocol Support**: Supports NFS, SMB, S3, and native POSIX clients, enabling multi-workload environments.
- **Future-Ready Design**: Purpose-built for modern workloads including AI/ML, HPC, analytics, media processing, and cloud-native applications.

Conclusion

The combination of LeilFS and OpenFlex Data24 4000 Series creates a powerful, next-generation storage solution that delivers hyperscale performance with enterprise-grade reliability. By leveraging NVMe-oF and distributed file system architecture, this solution enables organizations to efficiently handle massive data workloads with low latency, high throughput, and seamless scalability.

It eliminates traditional storage bottlenecks, reduces operational complexity, and provides a flexible foundation for modern data-intensive applications—making it an ideal choice for organizations aiming to build scalable, high-performance, and future-ready storage infrastructures.

- With this solution customers benefit from seamless scalability, expanding storage capacity without disrupting operations or application availability.
- It Improves cost efficiency and resource utilization through shared NVMe storage and composable infrastructure.
- Enhanced operational agility, allowing IT teams to dynamically allocate resources based on workload demands.
- Supports true disaggregation and composability, allowing independent scaling of compute and storage resources.
- Ensures high availability and data resilience through distributed architecture and redundancy mechanisms.
- Maximizes infrastructure utilization while reducing costs associated with overprovisioning.
- Provides a highly scalable, flexible, and efficient data fabric for modern enterprises.

5601 Great Oaks Parkway
San Jose, CA 95119, USA
www.westerndigital.com

August 2026 / Revision 01